# Meeting on 21st of April 2015

**Use cases**

**1)** HBP is developing an image service which manages all issues that providing multi-range solution of the use onto the data, searching of sub-volume of data, also be able to do arbitrary facing angles. This is a basic service for core processing like to register data and align data to to standard template brain spaces. The service is developed using Python, HDF5. Data container is currently plug into several web clients using desktop analysis software.

**2)** HBP plans to deploy analytical software that runs either on a single thread or multiple threads to analyse, for example, to extract neuron morphology, which needs a tracing algorithm for large collaborative data neutrons which is recently launched a world-wide efforts of collection of large neuron image sets. The aim is to use automatic reconstruct algorithm to trace these neutrons and extract the object that from the image sets which them be analysed by cluster algorithm. This may require to run ideal in parallel reconstruction algorithms and be able to compare the results.

HBP is looking for different type of repositories that not just create a collection of large images that have to be transferred. The intention is to active data services for accessing just sub-portion or resolution of that data wanted and ultimate be able to do analysis where that data sets are. The purpose is to leave the data in place.

**Computation Facilities:**

Imagine processing facilities include: image service that running operational in supercomputer centres; web clients operational for 2D/3D; desktop workstation platform with analysis program and interface to the service operations. Request for Core 15G RAM, high I/O storage as minimal configurations. They will collect information about standard use load.

**Image data size**: current image data sets are around 1-10 TB. With the quick growth of the imaging technology, the data collection can be grown quickly. E.g. in China it makes commercial for scanning. However, HBP does not expect their data grow will exceed the nowadays computing capabilities.

**I/O (file store to service)**:1GB/s under low testing condition

**Replications**:

Currently data are not replicated. However for optimising the performance, they will consider to replicate some key datasets to sites where lower latency is desirable. For example, there is a use case, where they are looking for to replicate large datasets from China to Europe.

**Data Access**:

Typically users will access data through the browsers. But they also have desktop clients with multi-resolution volume neurone analysis functions. There is a need to provide tool/service for data producers to upload data to the repository, high performance protocol for data transferring will be preferable to get those data to the repository effectively. There is a need to support all different types of data producers, e.g. to provide one default efficient transferring services, and a set of recommended tools/services. Currently, large data producers are those associated with supercomputer centres; others may have lab service or university server but don’t necessary have e-infrastructures to sustain the in a long-term, there are also use cases e.g., a group in China would like to make their data more widely available.

Data are open. But some data will need for AAI, e.g. some data producers want to limited their data access for collaborations. In general there is no identifiable individual human data.

Data access pattern: Depending of different use cases, arbitrary facing angle will need random data access, but many computation, e.g., for quick computed structure and quick computed projection are serial/sequential data access.

**Architecture**:

Centralised data archive within one data centre. The intention is to provide a long-term/permanent data access point.

**Dataflow**:

Data producers transfer data to the central data archive, data would be accessible without having to replicate the whole dataset but having the level of the details what they would be interested or the sub-portion that necessary for the analysis. Local data analysis workflow will need data injection at different stages. Raw data can be precompute and will be directly stored in the HDF5 files, but in general, many of them will be generate on the fly and not stored. e.g., users just rapidly navigating through the volume and search for a sub-volume of data and project them. However, in the second use case, there will be a need to capture immediate processing objects, since end user’s activities will be very exploratory, they may need to reuse the same results for later uses. Such processing results are not significant in volume, which is in MB level.

Data will be reduced during processing pipeline. There is few cases that processed data will be bigger than the raw datasets.

**Imagine processing**:

Currently, some image processing happen in the image service on the server sites, and the rest happen with web client in the browsers.

**Metadata**:

Some metadata are in HDF5 files referring to the ontology that operating on the ontology services. Other metadata can be searched by the knowledge graph.

**Data Identification**:

Current system has index to data and to metadata, searching facilities are provided allowing data discovery. Each dataset is associated with a global unique identifier and there are references (URIs) for multiple representations. For example, there is an entry for a unique dataset and its replica URIs, they link to a common GUID in the metadata system.

One details is the requirement for I/O performance which mentioned in previous conversations: which requests **1GB/s per node.**

**Requirements for Cloud platform**:

There are interesting/plan for CDMI development.

#

#

# Meeting on 13th of May 2015

Participants:

Sean Hill, HBP, **sean.hill**

Catherine Zwahlen, HBP, **cathzwah**

Lukasz Dutka, EGI Virtual Team on Federated data access , **lukaszdutka**

Enol Fernandez, EGI cloud expert, **enol.fernandez.egi**

Diego Scardaci, EGI user outreach, **dscardaci**

Yin Chen, EGI user outreach, **yin.chen.egi**

**Agenda**:

1. Use Cases step by step
2. EGI Testbed factors of the success
3. Agenda for Lisbon HBP session

**Parties:**

**Brain Research Facility**. There are large amounts of image stacks or volumetric data are produced daily at brain research sites around the world. This includes human brain imaging data in clinics, connectome data in research studies, whole brain imaging with light-sheet microscopy and tissue clearing methods or micro-optical sectioning techniques, two-photon imaging, array tomography, and electron beam microscopy.

**Active Repository Center.** The actual data from Brain Research Facility must be stored in data centers having respective storage solutions plus capacity for data processing.

**Metadata Brain Center.** Maintain the central HBP metadata repository. The repository is a kind of directory for all the available data in Active Repositories and it maintains as well LDAP of HBP members and OpenID interface for authentication.

**UC1: Brain Scan Creation**

**Actors**:

**Brain Researcher** working in a group/project performing brain scans. They work for brain research facilities around the world.

**Data Manager** working at Active Repository Center, responsible for maintaining data.

**Action:**

A Scientist is creating a brain scan, which is stored in a form of files. Then the files should be transferred to one of Active Repository Centers and register as well in the central metadata repository. Some metadata are included in the file but most of them are stored in JSON and XML file. The metadata are important for finding the right scan in the global metadata repository. Current metadata of scans are: resolution, species, size of the file, number, etc. They might be threaded as list of key/values.

Scans are stored in a form of: series of bitmaps,VTK (for 3d rendering), HDF5, TIFF/JPEG at origin,convert to HDF5 From the data structure point of view a single scan is always a directory of files, which should be registered in the central server.

→ For the operation of the image server, there is a single HDF5

there are can be a graph data model which links with many data sets then convert to HDF5.

**Current Solution:**

The images are transferred manually to Active Repository Centers where they will be processed later. In many cases FTP protocol is involved. Brain Researchers upload the scans to ftp server. In other cases Brain Researchers give access to the data on their own FTP servers to be downloaded by Data Manager, or in some worst case of the largest data sets scenarios Brain Researches send hard drives with data to be uploaded.

**Problems to be solved related to UC1:**

* Data flow from Brain Research Facilities to Active Repository Centers.
* Selection of Active Repository Centers to which center the scan should be delivered. There is plan to build multiple processing centers in the world, ideally one per country or more.
* Capacity management, how to maintain grants for storage for different groups of scientists.
* How to replicate data between Active Repository Centers.--> replication is needed when latency is too high eg. China or in some cases west coast in the US

 when replicating data, the location of the replicas will be stored in the metadata.

* Metadata management support

**Output:**

* Sets of directories containing brain scans described by some metadata files in JSON or XML formats

**UC2: Remote interactive multiresolution visualization of large volumetric datasets**

**Actors**:

**Brain Researcher** there are brain researchers interested in navigating through existing brain scans.

**Actions:**

Brain Researchers navigate around the brain image using web browser. To make that possible first the Brain Researchers must find a respective scan using central metadata server - this is possible thanks to Metadata Brain Center. Then having direct link to Active Repository Center Brain Researcher is able to dive virtually into the brain thanks to the WebGL standard (--> not hard requirement, graphics acc. not required). But to make that possible Active Repository Center needs to deploy interface which produces WebGL data based on the actual registered image scans. The interface processes data using POSIX api and produces singigicanlty small amout of data in form of HTML+WebGL.

The process of generation WebGL data is IO intensive. The experiments showed that one physical server is able to handle max 10 simultaneous viewers. Expected number of simultaneous users still unknown, want to understand how to scale.

→ interactivity only relevant in the first case study, not the second

→ dynamic load balancing of clients in case of many users accessing the same data set of interest, but this should not be a problem in the short term (2 years)

**Current location of the data:**

CINECA, Juelich, Oslo, SuperComputer Center in Switzerland, UPM Spain

Designated repository per geographical location (the closest).

Current size of data collections is: xxx

**Problems to be solved related to UC2:**

* how to store the data, keep them available through high throughput POSIX interface but still having federated management functionality
* access control to data integrated with AAI based on OpenID (--> to start with AAI can be avoided with open data where auth/authz are not necessary)
* how to maintain data space (space allocation, cpu utilization etc.)
* how to easily deploy Active Repository Centers to make them as many as possible
* how to distribute software for active repository centers. In other words releasing new software for image navigation should lead to updating gracefully all the Active Repository Centers (--> there will be many data producers around the world, intention is to provide easy to use means for them to upload the data)
* if the cloud will be a solution what would be the cost model. (depend on which cloud and which cloud model. )
* Storage QoS taking into account UC3, which might degrade storage performance for UC2 keeping in mind that UC2 is more interactive and UC3 is more batch processing.

how to deal with pick time access

when request high availability, may need prevision data generation and processing

need to understand the user access patterns. But HBP in the next 2 years may not have the problem.

**UC3: Feature extraction and analysis of large volumetric datasets**

**Actors:**

**Neuroscientist** - an actor trying to generate new data based on existing brain scans and register new data into the central metadata server. (Brain researcher is more general term. They have similar access behaviours and access right)

**Actions:**

From the technical point of view in this usecase neuroscientists process data directly from the repositories and generates new data which should be registered in the global metadata server.

**Detailed Description:**

This use case would entail deploying Vaa3D (www.vaa3d.org) as an additional service to the active repository described in Use Case I. Vaa3D is open source and provides a plugin architecture into which any type of neuron reconstruction algorithm can be adapted. The second use case would require additional computational resources (and could benefit from multithreaded and parallel compute resources) for the reconstruction process.

In this use case, a neuroscientist user would provide via a web service input parameters to a Vaa3D instance which would trace any recognized neuron structures using a selected algorithm. The output file would be returned via the webservice.

Output is small and can be transferred on REST interface.

No specific hw needed, 4 cores

**Problems:**

* How to give neuroscientists access to the actual data and provide them in the same time possibility to generate new data. The generated data needs to be stored somewhere keeping in mind limited access and yet registered in the central metadata.
* How to efficiently process the data on cloud or grid infrastructure

data is small

**Use case requirements:**

* The active repository developed for UC2.
* The additional deployment of Vaa3D ([www.vaa3d.com](http://www.vaa3d.com) ) adapted for use with BBIC. A beta version of this is currently available. The REST API may need development.
* A multiprocessor compute node with high speed access to the storage device.
* Additional datasets including image stacks/volumes of clearly labeled single or multiple neurons - provided by HBP, Allen Institute and others. (data can be processed at each place, not issue for migration)
* Scaleable computing environment.

**Output:**

Extracted data objects …. keeping full control of the access rights to data owners.

**UC4: Publication and citation of data**

**Actors:**

 **Data Owner:** neuroscientists having administrative access to the data

**Actions:**

Data Owners should be able to generate persistent citable links to data. It should be work as well with DOI. However by generating citicable links data owner will need to take some responsibility to keep that data in the longer perspective. Data owner can define access level to the cited data including anonymous level, by that granting access to anyone having the link. In other cases access control on the cited data should be still possible.

Access to the "cited data" should be monitored to gain ability to generate some data access statistics.

In this use case any object could be cited including: original data sets, subvolumes, extracted object, etc.

A challenging issue in this use case might be ability to cite data being an output of a processing services. For instance an actor wants to cite a part of brain scan which is accessible via image navigation services but not necessarily to copy that data.

->policy about deletion of data

**Problems:**

* integration of high throughput data repositories with concepts of persistence and citation of data
* limited access to the features of long term data preservations
* forking of data between data owners to overtake responsibilities of long data preservation
* citing data that have no physical representation but being output of other processing services without storing that. However to achieve long term persistence it might be required to copy the cited data to avoid future problems when the processing service changes and might generate different outputs.

**UC5: Management of Access Control Rights**

**Actors:**

**Data Owner:** neuroscientists having administrative access to the data

**Actions:**

Human Brain Project maintains its own LDAP repository of users and groups. Based on that there is an OpenID provider which supports identification of the users.

There must be an easy way to maintain ACL rights based on groups membership. Each scan or data object could have an individual ACL rights maintained by the users having respective permissions.

 ->at the moment is only accessible by authenticated users, including the user processing off-line

->repository interests decide the access controls

**Requirements:**

* System must be compatible with OpenID
* ACL

## REQUIREMENTS FOR Infrastructure Providers:

* **PERFORMANCE**. Being able to work at throughput 1GB/s per node to support 10 users (for image service nodes, throughput between the data server and the storage)
* **STORAGE CAPABILITIES**
	+ - files at size up to 10TB
		- IOPS requirements (I/O operations per second) ??
		- Posix access to data
* **CO-LOCATION**. cpu close to data

**Metrics for success of EGI tests:**

* simplicity of bringing up another active repository
* simplicity of the process transferring data to the active repository site (UC1)
* flexibility of accessing the same data by multiple scientists including intra-groups access (UC3)
* decentralization of resource management. There are many collaborating groups but they remain independent. The system should be flexible to allow independently gain resources by those groups and still provide some integration level.
* simplicity of the access control (UC5) ?
* scalability of processing through load distribution/brokering, it should be easy (maybe automatic), instantiation of new processing units based on the traffic to active repositories. (UC2)
* how difficult distribution/upgrading of the application software onto the environment will be. Docker approach is under HBP investigation now.
* support for persistent identification (DOI and persistence life cycle approaches) (UC4) (->in long-term will be important)
* process of migration data between active repositories if needed, that might be limited (->next step requirement)

Involvements in the TestBed: Open Science, ...

Internal release in a month, publication in 1 year

Meeting on 21st of May 2015 (EGI Conference)

Participants: to be added

Sean Hill introduced the 2 Human Brain (HB) use cases

Lukas Dukta presentation:

* EGI FedCloud intro
* Federated Solutions for data management current available:
	+ CDMI (GRNET)
	+ dCache (DESY)
	+ DinaFed (CERN)
	+ FTS (CERN)
	+ iRODS (CNRS)
	+ Onedata (CYFRONET)
* Deployment
	+ unique metadata catalogue with information about the location of the images
	+ more active repository instances in the world, multiple repositories with no intention to ship large amounts of data, the amount of data from server to client is limited, no link between the data centres (working independently)
	+ the computation is performed in the active repository where the image is available (close the data) and closest the user.
	+ image replications for improve interaction is considered (e.g. image stored in China and replicated in Europe).
	+ federation between the 4 HPC centres, with PRACE dedicated links (some data research facilities already connected in the last mile at 10 Gbps), but in some cases the network could be not as good
	+ 5 use cases identified (see descriptions above)
	+ the active repository should be easily replicated in different geographical locations.

Tiziana Ferrari (TF): HB already tested iRODS and it seemed not suitable for the 2 use cases.

Sean Hill (SH): AAI interoperability, type of interfaces offered, some instability (may have improved)

Giuseppe Fiameni (GF): the issue with iRODS is the access. It is not scale well with many files.

GF: the HPC platform is connected with PRACE. Dedicated connection (1GB network) with some brain research facilities (resource centers) to optimise the last mile. Gridftp is used.

ACTION (Tiziana). Contact Steve Tycke / GLOBUS to see if globus online could a useful service

Action (SH): provide the XML file describing a scan.